Title	暗号における組合せ構造について(離散数理モデルにおける最適組合せ構造)
Author(s)	岡本, 栄司
Citation	数理解析研究所講究録 (1993), 820: 54-58
Issue Date	1993-02
URL	http://hdl.handle.net/2433/83176
Right	
Type	Departmental Bulletin Paper
Textversion	publisher

数理解析研究所講究録
第820巻 1993年 54-58

暗号における組合わせ構造について

岡本栄司
北陸先端科学技術大学院大学
情報科学研究科
e-mail: okamoto@jaist-east.ac.jp

1 はじめに

暗号は情報発信者が受信者のみに意味がわかるように情報を伝える技術であり、認証は当事者・情報の正当性を保証する技術である。これらの暗号／認証においてベースとなる数学的な理論には、情報理論、初等整数論、計算量理論、組合わせ理論などがある。それらは次のような役割を担っている。

- 情報理論
  - 安全性保証（計算機パワー無限）
  - 情報帯域の下限
- 初等整数論
  - 公開鍵暗号系、ディジタル署名系構成（剰余演算）
  - 解読（素因数分解、離散対数）
- 計算量理論
  - 安全性保証（計算複雑度、計算機パワー有限）
- 組合わせ理論

組合わせ理論は暗号／認証において多く用いられているが、中でも、暗号アルゴリズム、認証、鍵配送／管理の各部門で効果的に用いられている。ここでは、トピック的にそれらの例を示し、その有効性を探る。

2 暗号アルゴリズム

2.1 誤り伝搬

通信暗号では、暗号化情報の伝送中に誤りが生じると、たとえ1ビットエラーでも受信側復号中にそのエラーが広がる（エラー伝搬）。これは、衛星通信などへの適用を考えると好ましくない。そこで、まず、エラー伝搬を起こさない暗号アルゴリズムはどういうものかを探る。

定理 1 エラー伝搬を起こさないための全単射暗号の条件は、暗号変換が

$$y = Px \oplus a \tag{1}$$

で与えられることである。ただし、x は平文ベクトル、yは暗号文ベクトル、Pは転置行列、a は定数ベクトル、$\oplus$ はビット毎の排他的論理和である。

(証明) 復号時に誤り伝搬を起こさない暗号変換を fとし、平文ベクトルと暗号文ベクトルをそれぞれ n 次元ベクトル x, yとおく:

$$y = f(x) \tag{2}$$

ここで、e_iを成分 i のみが 1 で他は全て 0 の単位ベクトルとおく。e_iは 0(零ベクトル) からハミング距離 1 だけ離れているので、$f^{-1}(0)$ と $f^{-1}(e_i)$ もハミング距離 1 だけ離れているはずである。従って、

$$f^{-1}(e_i) = f^{-1}(0) \oplus e_{\sigma(i)} \tag{3}$$

となる。

次に、これを用いて、一般に $y = \sum_{i=1}^{n} y_i e_i$のとき

$$f^{-1}(y) = a \oplus \sum_{i=1}^{n} y_i e_{\sigma(i)} \tag{4}$$

を y の重みに関する数学的帰納法にて示す。ここで、$a = f^{-1}(0)$ とおいた。重み 1 のときは既に示してある。重みが k 以下のときは正しいとして、y の重みが k+1 とする。y の最初の 2 つの非零成分を抜き出して

$$y = e_{i_1} \oplus e_{i_2} \oplus z \tag{5}$$

とする。数学的帰納法の仮定から、

$$f^{-1}(z) = a \oplus \sum_{i \neq i_1, i_2} y_i e_{\sigma(i)} \tag{6}$$

$$f^{-1}(e_{i_1} \oplus z) = a \oplus \sum_{i \neq i_1} y_i e_{\sigma(i)} \tag{7}$$

$$f^{-1}(e_{i_2} \oplus z) = a \oplus \sum_{i \neq i_2} y_i e_{\sigma(i)} \tag{8}$$

が成り立つ。yと $e_{i_1} \oplus z$、あるいは $e_{i_2} \oplus z$との間は共にハミング距離 1 なので、yに対する平文ベクトルは $a \oplus \sum_{i \neq i_1, i_2} y_i e_{\sigma(i)}$か $a \oplus \sum_{i=1}^{n} y_i e_{\sigma(i)}$のいづれかしかない。しかし、前者は z に対応する平文ベクトルなので、後者が yに対する平文ベクトルである。

ここで、式 (3) の右辺を x とおくと

$$\sum_{i=1}^{n} y_i e_{\sigma(i)} = a \oplus x = \sum_{i=1}^{n} (a_i \oplus x_i) \tag{9}$$

から、

$$y_i = a_{\sigma(i)} \oplus x\sigma(i) \tag{10}$$

となる。従って、

$$f(x) = y = \sum_{i=1}^{n} y_i e_i = \sum_{i=1}^{n} (a_{\sigma(i)} \oplus x_{\sigma(i)}) e_i = p \oplus \sum_{i=1}^{n} x_{\sigma(i)} e_i \tag{11}$$

が得られ、定理が証明できた。(証明終り)

誤り伝搬が少し起きる場合の議論も行なわれているが、完全な構造解明には至っていない [MO91]。

2.2 秘密情報分散法 [Ok92]

重要な秘密情報を保管するのに、分散化する方法がある。(k,n) 閾値法は n 人中 k 人集まれば元の情報を復元でき、$k-1$ 人以下では元の情報の1ビットすら得られないという方法である。

この (k,n) 閾値法は、実用を目指して拡張されている。最も一般的な拡張は、次の単調性である。

定理 2 秘密情報分散可能となる条件は、単調性：

$$X \subseteq Y, X \in A \Rightarrow Y \in A \tag{12}$$

である。ここで、A は元秘密情報復元可能な人の集まり、X、Y は人の集まりである。

（証明）秘密情報分散法があれば、単調であるのは明らかである。十分条件を示す。S を $GF(p)$ における秘密情報とする。A が単調ならば、A の任意の極小グループに対して、そこに属すメンバー $i_1, i_2, \cdots$ に

$$S = a_{i_1} + a_{i_2} + \cdots (\mathrm{mod}\, p) \tag{13}$$

となる $a_{i_1}, a_{i_2}, \cdots$ をランダムに生成して、渡す。これを各極小グループ毎に行なう。これは、秘密情報分散法になっている。（証明終り）

この一般化秘密情報分散法の欠点は、各人が持つべき分割情報が n の指数関数的に比例することである。そこで、式（12）の条件を多少変えて、各人が持つ分割情報が元の情報と同じ大きさとなる理想分散法の研究されている。しかし、上記定理2を見ればわかるように、秘密情報 S は、極小グループ毎に互いに独立な情報とすることができる。このようにしたときには、秘密情報量／分散情報量は、1まで上げられる。なお、1以上にすることは実用上余り意味がない。

3 認証

認証においては、ディジタル署名、メッセージ認証、ユーザ認証のいずれをとっても、ハッシュ関数を必要とする。暗号分野でいうハッシュ関数はいわゆる一方向性ハッシュ関数 h で、

$$h(x) = h(y) \tag{14}$$

を満たす（衝突する）x, y を容易に見い出せないことをいう。

定理 3 $h : X \mapsto Y$ が全射とする。$N = |Y|$ ならば、X から $\sqrt{N}$ 個の元をもってくると、その中に衝突する入力ペアの存在する可能性は高い。実際には X から $1.18\sqrt{N}$ の元をもってくると、衝突する確率は $1/2$ 以上になる。

（証明）Y から重複を許して k 個の $a_1, a_2, \cdots, a_k$ を選んだとき、これらが全て異なる場合は、a_2 が a_1 と異なり、かつ a_3 が a_1, a_2 と異なり、$\cdots$、かつ a_k が $a_1, a_2, \cdots, a_{k-1}$ と異なる場合である。したがって、求める確率 P は、

$$P = 1 - \frac{n-1}{n}\frac{n-2}{n}\cdots\frac{n-k+1}{n} = 1 - (1-\frac{1}{n})(1-\frac{2}{n})\cdots(1-\frac{k-1}{n}) \tag{15}$$

で与えられる。これを自然対数を用いて近似すると、$x = \frac{k-1}{n}$ として、

$$\frac{1}{n}\log(1-P) = \frac{1}{n}\sum_{j=0}^{k-1}\log(1-\frac{j}{n})$$

$$\simeq \int_{1-x}^{1}\log t dt = -x - (1-x)\log(1-x) = -\frac{x^2}{2} - \frac{x^3}{3} - \cdots \tag{16}$$

を得る。$P = \frac{1}{2}$となるkを求めるには、上式から、解xが小さいことを利用して解けばよく、

$$\frac{x^2}{2} \simeq \frac{1}{n} \log 2 \tag{17}$$

$$k \simeq \sqrt{2 \log 2n} = 1.18\sqrt{n} \tag{18}$$

が得られる。$k = \sqrt{n}$のときは、確率は$P = 0.4$になる。（証明終り）

上記定理3はいわゆるバースディパラドックスである。これによれば、認証の標準として定められているDESを用いた認証法は、$N = 2^{32}$を採用しているので、弱い。実際、$\sqrt{N} = 2^{16} = 65,536$個の入力について衝突を検査するのは容易である。

なお、バースディパラドックスの定式化にはもう一つある。クラスの中で、誕生日の一致する男女のペアが存在する確率$\frac{1}{2}$を越すためには、クラスの大きさはどのくらい必要かというもので、実際にはこのタイプが解読に用いられる。スケールが大きくなると、どちらの定式化でも同じ結果となる[Ni90]。

4 暗号鍵

暗号鍵は暗号変換のパラメータと考えられる。暗号強度が高いためには、暗号鍵数は多い必要がある。しかし、暗号鍵の個数だけをいくら多くても、暗号文を正しい鍵以外の鍵で復号して元の平文に近い文が出てくるようではまずい。この場合、似た効果しかもたらさない鍵は「同じ」鍵として扱うべきである。このとき、「異なる」鍵の個数を**実質鍵数**[Ok88]といい、この実質鍵数が多い必要がある。

定理 4 鍵空間において、半径ϵの球に属す鍵の個数をQ_ϵとする。このとき、実質鍵数は

$$SNK = N_K / Q_\epsilon \tag{19}$$

で与えられる。ここで、ϵは「同一」とみなすか否かの境界を示す閾値であり、N_Kは全鍵数である。

（証明）鍵空間における全鍵数をN_kとすると、任意に選んだ鍵Kがある鍵K_0を中心として半径ϵの球に属する確率は、$\frac{Q_\epsilon}{N_k}$で与えられる。一方、一般的に、互いに異なる碁石がSNK個存在したとき、任意に選んだ碁石がある特定の碁石に一致する確率は$\frac{1}{SNK}$である。そこで、これらの確率を等号で結べば、式（19）を得る。（証明終り）

ここで、ビット反転率[Na80]に基づいて、乱数加算型の簡単な暗号$y = x \oplus K$の実質鍵数を具体的に求めてみよう。長さnの二つのビット系列x_1, x_2のビット反転率$d(x_1, x_2)$を

$$d(x_1, x_2) = \frac{\mathrm{Hamming}(x_1, x_2)}{n} \tag{20}$$

で定義する。ここで、$\mathrm{Hamming}(x_1, x_2)$は$x_1, x_2$の間のハミング距離である。ただし、反転率は0.5のときが最も離れており、1のときは実は距離0と同じとみなすべきである。そこで、修正したビット反転率$r(x_1, x_2)$:

$$r(x_1, x_2) = \frac{1}{2} - \left| \frac{1}{2} - \frac{\mathrm{Hamming}(x_1, x_2)}{n} \right| \tag{21}$$

を用いる。

さて、暗号変換を

$$y = x \oplus K \tag{22}$$

とする。ここで、Kは任意のn次元バイナリベクトルをとり得るものとする。このとき、ある鍵Kから半径ϵの球に属する鍵の個数Q_ϵは、

$$Q_\epsilon = |\{x : \text{Hamming}(K,x) \leq n\epsilon \text{ または } \text{Hamming}(K,x) \geq n(1-\epsilon)\}|$$

$$= |\{x : \text{weight}(x) \leq n\epsilon \text{ または } \text{weight}(x) \geq n(1-\epsilon)\}| \tag{23}$$

で与えられる（weight(x)は、xの中の1の個数を示す）。これは、誤差関数を用いて近似できる。すなわち、

$$\sum_{i=0}^{k} \binom{n}{i} \simeq 1 - \text{erf}\left(\frac{k-np}{\sqrt{npq}}\right) \tag{24}$$

$$p+q=1.$$

ここで、

$$\text{erf}(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \simeq \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}}. \tag{25}$$

この誤差関数を用いると、実質鍵数は

$$SNK = \frac{2^n}{Q_\epsilon} = \frac{1}{2\,\text{erf}(1-2\epsilon)\sqrt{n}} = \sqrt{\frac{n\pi}{2}}(1-2\epsilon)e^{\frac{(1-2\epsilon)^2 n}{2}} \tag{26}$$

となる。

オーダーを見るために、2を底とする対数をとると、主要項は

$$\frac{(1-2\epsilon)^2 n}{2} \log_2 \epsilon \tag{27}$$

であり、見かけ上の鍵ビット長nが減っていることがわかる。

References

[MO91] 宮野, 岡本: 暗号における誤り波及に関するグラフ理論的一考察, 電子情報通信学会情報セキュリティ研究技術報告 ISEC90-51, pp.47-53, 1991

[Na80] 中村: 自己同期型簡易暗号方式に関する一考察, 第3回情報理論とその応用研究会予稿集, pp.371-377, 1980

[Ni90] Nishimura: Probability to meet in the middle, Journal of Cryptography, Vol.2, pp.13-22, 1990

[Ok88] Okamoto: Substantial number of cryptographic keys and its application to encryption designs, Proc. of Eurocrypt'88, pp.361-373, 1988

[Ok92] 岡本, 秘密情報分散管理方式の構造について, SCIS'92, 1992